DECOMPOSITION OF NON-NULL WILKS' U AND MULTIVARIATE BETA

by

Somesh Das Gupta*

University of Minnesota

Technical Report No. 164

November 1971

University of Minnesota

Minneapolis, Minnesota

*This research was supported by United States Army Grant DA-ARO-D-31-124-70-G102.

DECOMPOSITION OF NON-NULL WILKS' U AND MULTIVARIATE BETA

1. Introduction.

A simple proof is given for Bartlett's decomposition of Wilks' U in the non-null case. Using this result a decomposition theorem is proved for multivariate B distribution in the non-null case. This generalizes the corresponding result for the univariate B distribution. Furthermore, some properties of the non-null multivariate beta distribution are derived.

2. Bartlett's Decomposition in the Non-Null Case.

Let $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n$ be independently distributed random $p \times 1$ vectors such that $X_\alpha \sim N_p(\mu_\alpha, \Sigma)$ and $\mu_\alpha = 0$ for $\alpha = r+1, \ldots, n$. Define

$$U = \left| \sum_{\alpha=r+1}^{n} X_\alpha X_\alpha' \right| / \left| \sum_{\alpha=1}^{n} X_\alpha X_\alpha' \right|.$$

Assume $n - r \geq p$. Let

$$M' = [\underset{\sim}{\mu}_1 \;\cdots\; \underset{\sim}{\mu}_r].$$

We denote the distribution of U by

$$U_p(n-r,\ r\ ;\ \mathrm{ch}(\Sigma^{-1}M'M)),$$

where ch indicates the characteristic roots.

For $p = 1$, $U_1(n-r,\ r;\ \lambda)$ = noncentral beta $(\frac{n-r}{2}, \frac{r}{2}; \lambda)$. For $r = 1$, $U_p(n-1,\ 1;\ \lambda)$ = noncentral beta $(\frac{n-p}{2}, \frac{p}{2}; \lambda)$, where λ is the only nonzero latent root of $\Sigma^{-1}M'M$. These special cases are well-known.

For $1 \leq q < p$, let

$$\begin{bmatrix} \underset{\sim}{X}'_1 \\ \vdots \\ \underset{\sim}{X}'_n \end{bmatrix} = \underset{n\times p}{X} = \begin{array}{cc} \left[\begin{array}{c|c} X_{11} & X_{12} \\ \hline X_{21} & X_{22} \end{array}\right] & \begin{array}{l} r \\ n-r \end{array} \\ \begin{array}{cc} q & (p-q) \end{array} & \end{array} = \begin{array}{cc} [X^{(1)} & X^{(2)}] \\ q & (p-q) \end{array}$$

$$S = \sum_{\alpha=r+1}^{n} X_\alpha X'_\alpha = \begin{array}{cc} \left[\begin{array}{c|c} X'_{21}\ X_{21} & X'_{21}\ X_{22} \\ \hline X'_{22}\ X_{21} & X'_{22}\ X_{22} \end{array}\right] & \begin{array}{l} q \\ (p-q) \end{array} \\ \begin{array}{cc} q & (p-q) \end{array} & \end{array} = \left[\begin{array}{c|c} S_{11} & S_{12} \\ \hline S_{21} & S_{22} \end{array}\right]$$

$$T = \sum_{\alpha=1}^{n} X_\alpha X'_\alpha = X'X = \left[\begin{array}{c|c} X'_{21}\ X_{21} + X'_{11}\ X_{11} & X'_{21}\ X_{22} + X'_{11}\ X_{12} \\ \hline X'_{22}\ X_{21} + X'_{12}\ X_{11} & X'_{22}\ X_{22} + X'_{12}\ X_{12} \end{array}\right] = \left[\begin{array}{c|c} T_{11} & T_{12} \\ \hline T_{21} & T_{22} \end{array}\right].$$

$$S_{22\cdot 1} = S_{22} - S_{21}S_{11}^{-1}S_{12}\ , \quad T_{22\cdot 1} = T_{22} - T_{21}T_{11}^{-1}T_{12}\ .$$

Note now

$$|S|\ /\ |T| = [|S_{11}|\ /\ |T_{11}|]\ [|S_{22\cdot 1}|\ /\ |T_{22\cdot 1}|].$$

We shall consider first the conditional distribution of $X^{(2)}$, given $X^{(1)}$. Conditionally, the rows of $X^{(2)}$ are independent $N_{p-q}(\cdot, \Sigma_{22\cdot 1})$ where

$$\Sigma_{22\cdot 1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\ , \quad \Sigma = \begin{array}{cc} \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right] & \begin{array}{l} q \\ (p-q) \end{array} \\ \begin{array}{cc} q & (p-q) \end{array} & \end{array}$$

$$E[X^{(2)}|X^{(1)}] = \underset{n\ \times\ (r+q)}{\left[\begin{array}{c|c} I_r & X_{11} \\ \hline 0 & X_{21} \end{array}\right]} \quad \underset{(r+q)\ \times\ (p-q)}{\begin{bmatrix} \alpha' \\ \\ \beta' \end{bmatrix}}$$

where $\beta = \Sigma_{21}\Sigma_{11}^{-1}$, $\alpha = M'_2 - \beta M'_1$, $M = \begin{array}{cc} [M_1 & M_2] \\ q & (p-q) \end{array}$. Using this conditional model, the maximum likelihood estimates of α and β are

$$\hat{\alpha} = X_{12}' - \hat{\beta}\, X_{11}' \;,\; \hat{\beta} = (X_{22}'\, X_{21})(X_{21}'\, X_{21})^{-1}.$$

The residual (error) matrix is $S_{22\cdot1}$. Under the restriction $\alpha = 0$, the residual error matrix is $T_{22\cdot1}$. It can be seen easily that

$$T_{22\cdot1} - S_{22\cdot1} = \hat{\alpha}\, G^{-1}\, \hat{\alpha}'$$

where

$$G = I_r + X_{11}(X_{21}'\, X_{21})^{-1} X_{11}' \;.$$

It follows from the standard results in MANOVA that, given $X^{(1)}$

$$S_{22\cdot1} \sim W_{p-q}(n-r-q,\, \Sigma_{22\cdot1})$$

$$T_{22\cdot1} - S_{22\cdot1} \sim W_{p-q}^{*}(r,\, \Sigma_{22\cdot1},\, \alpha\, G^{-1}\alpha)$$

and $S_{22\cdot1}$, $T_{22\cdot1} - S_{22\cdot1}$ are independent. In the above, W and W^* denote the central Wishart and the noncentral Wishart distributions. Hence, given $X^{(1)}$

$$\frac{|S_{22\cdot1}|}{|T_{22\cdot1}|} \sim U_{p-q}(n-r-q,\, r\, ;\, ch[\Sigma_{22\cdot1}^{-1}(\alpha\, G^{-1}\alpha')]).$$

Note also that

$$\frac{|S_{11}|}{|T_{11}|} \sim U_q(n-r,\, r\, ;\, ch[\Sigma_{11}^{-1} M_1' M_1]).$$

Let $\Lambda = |S| \,/\, |T|$, $\Lambda_1 = |S_{11}| \,/\, |T_{11}|$, $\Lambda_{2\cdot1} = |S_{22\cdot1}| \,/\, |T_{22\cdot1}|$. Then $\Lambda = \Lambda_1 \Lambda_{2\cdot1}$, and

$$\Lambda \sim U_p(n-r,\, r\, ;\, ch^*(\Delta_p))$$
$$\Lambda_1 = |G^{-1}| \sim U_q(n-r,\, r\, ;\, ch^*(\Delta_q))$$
$$\Lambda_{2\cdot1} \text{ given } G \sim U_{p-q}(n-r-q,\, r\, ;\, ch^*[G^{-1}(\Delta_p - \Delta_q)])$$

where ch^* denotes the non-zero characteristic roots, and Δ_p and Δ_q are defined as follows:

$$\Delta_p = M\Sigma^{-1}M', \quad \Delta_q = M_1\Sigma_{11}^{-1}M_1' .$$

Note that

$$\Delta_p - \Delta_q = (M_2 - M_1\Sigma_{11}^{-1}\Sigma_{12})\Sigma_{22\cdot1}^{-1}(M_2 - M_1\Sigma_{11}^{-1}\Sigma_{12})'.$$

The standard Bartlett's decomposition of U in the null case follows easily from the above results.

3. Decomposition and Properties of the Noncentral Multivariate Beta Distribution.

Let us now write

$$X = \begin{bmatrix} X'_{(1)} \\ X'_{(2)} \end{bmatrix} \begin{matrix} r \\ n-r \end{matrix} .$$

Define

$$\underset{r\times r}{B} = I_r + X'_{(1)}[X_{(2)}X'_{(2)}]^{-1}X_{(1)} .$$

We shall say B is distributed as the noncentral multivariate beta distribution

$$B_r(n-r,\ p\ ;\ M\Sigma M').$$

Let

$$B_1 = X_{11}(X'_{21}X_{21})^{-1}X'_{11} + I_r ,$$

i.e., B_1 is defined as B except for using the first q components. It is easy to see that

$$B - B_1 = [X_{12} - X_{11}(X'_{21}X_{21})^{-1}(X'_{21}X_{22})]S_{22\cdot1}^{-1}[X_{12} - X_{11}(X'_{21}X_{21})^{-1}(X'_{21}X_{22})]' .$$

Given $X^{(1)}$,

(a) $G^{-\frac{1}{2}}\hat{\alpha}'$, $S_{22\cdot 1}$ are independently distributed,

(b) $S_{22\cdot 1} \sim W_{p-q}(n-r-q, \Sigma_{22\cdot 1})$,

(c) the rows of $G^{-\frac{1}{2}}\hat{\alpha}'$ are independent $N_{p-q}(\cdot, \Sigma_{22\cdot 1})$ and the conditional expectation of $G^{-\frac{1}{2}}\hat{\alpha}'$ is $G^{-\frac{1}{2}}\alpha'$.

Hence

$$I_r + G^{-\frac{1}{2}}\hat{\alpha}'S_{22\cdot 1}^{-1}\hat{\alpha}G^{-\frac{1}{2}} \sim B_r(n-r-q, p-q ; G^{-\frac{1}{2}}\alpha'\Sigma_{22\cdot 1}^{-1}\alpha G^{-\frac{1}{2}}).$$

Define

$$B_{2\cdot 1} = B_1^{-\frac{1}{2}}(B-B_1)B_1^{-\frac{1}{2}} + I_r .$$

Finally, we have the following results:

$$B = B_1^{\frac{1}{2}} B_{2\cdot 1} B_1^{\frac{1}{2}}$$

$$B \sim B_r(n-r, p ; \Delta_p)$$

$$B_1 \sim B_r(n-r, q ; \Delta_q)$$

$$B_{2\cdot 1}, \text{ given } B_1 \sim B_r(n-r-q, p-q ; B_1^{-\frac{1}{2}}(\Delta_p - \Delta_q)B_1^{-\frac{1}{2}}).$$

This result is interesting since it generalizes the corresponding result for univariate non-central beta distribution. To see this, we proceed as follows.

For $r = 1$, $m = n - r$

$$1/B \sim \text{noncentral beta } (m-p+1, p ; \Delta_p)$$

$$1/B_1 \sim \text{noncentral beta } (m-q+1, q ; \Delta_q)$$

$$1/B_{2\cdot 1} \text{ , given } B_{11} \sim \text{noncentral beta } (m-p+1, p-q ; (\Delta_p - \Delta_q) \frac{1}{B_1})$$

and

$$1/B = (1/B_1) \cdot (1/B_{2\cdot 1}).$$

Let us consider some other properties of the B_r distribution. First note that, if $B \sim B_r(m, p ; \Delta)$ then $|B|^{-1} \sim U_p(m, r ; \Delta)$. Let b_{11} be the first diagonal element of B. Then

$$b_{11} \sim B_1(m, p ; \delta_{11}) = 1/\text{noncentral beta } (m-p+1, p ; \delta_{11}),$$

where δ_{11} is the first diagonal element of Δ.

Let $1 \leq s < r$.

$$\underset{r \times r}{B} = \begin{array}{cc} \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right] & \begin{array}{c} s \\ r-s \end{array} \\ \begin{array}{cc} s \quad & \quad r-s \end{array} & \end{array} , \quad \Delta = \begin{array}{cc} \left[\begin{array}{c|c} \Delta_{11} & \Delta_{12} \\ \hline \Delta_{21} & \Delta_{22} \end{array}\right] & \begin{array}{c} s \\ r-s \end{array} \\ \begin{array}{cc} s \quad & \quad r-s \end{array} & \end{array}$$

$$B_{22\cdot 1} = B_{22} - B_{21}B_{11}^{-1}B_{12} .$$

Let

$$\underset{r \times p}{X'_{(1)}} = \begin{array}{cc} \begin{bmatrix} U'_1 \\ U'_2 \end{bmatrix} & \begin{array}{c} s \\ r-s \end{array} \end{array} .$$

Recall that

$$B = I_r + X'_{(1)}S^{-1}X_{(1)} .$$

Then

$$B_{11} = I_s + U'_1S^{-1}U_1 \sim B_s(m, p ; \Delta_{11})$$

$$B_{22\cdot 1} = I_{r-s} + U'_2(S + U_1U'_1)^{-1}U_2 .$$

Suppose now $\Delta_{11} = 0$, i.e., $E(U_1) = 0$. Then, clearly B_{11} and $B_{22\cdot 1}$ are independent, and

$$B_{22\cdot 1} \sim B_{r-s}(m+s, p ; \Delta_{22}).$$

This gives another decomposition of B. Also, it is easy to see that for any non-null vector $\underset{\sim}{a} : r \times 1$

$$\frac{a'Ba}{a'a} \sim B_1(m, p ; a'\Delta a).$$

Consider now

$$\underset{\sim}{a}'B^{-1}\underset{\sim}{a}$$

when

$$B \sim B_r(m, p ; 0) .$$

It is clear that for any orthogonal matrix $L : r \times r$, $LBL' \sim B_r(m, p ; 0)$.

Hence

$$\frac{a'a}{a'B^{-1}a} \sim \frac{1}{b^{11}} .$$

In the previous result, let $s = r - 1$. Then

$$\frac{1}{b^{11}} \sim B_{22\cdot1} \sim B_1(m+r-1, p ; 0).$$

Finally, note that if $p = 1$, $|B|^{-1} \sim$ noncentral beta $(m, r ; tr(\Delta))$.

Remark 1.

All the above results for decomposition of U and B can easily be generalized to more than two factors.

Remark 2.

For the standard results used in this note, one may refer to Anderson [1] or Rao [7]. Bartlett's decomposition of non-central U was previously obtained by Katti [3] when rank(M) = 1 and by Asoh and Okamoto [2] when rank(M) = 1 or 2. A direct proof is given here for the general case. For multivariate beta distribution (central) one may refer to Olkin and Rubin [6] , Mitra [5], and Khatri [4]. In these papers, some other versions of the multivariate beta distributions are also considered.

References.

[1] Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, New York.

[2] Asoh, Y. H. and Okamoto, M. (1964). A note on the non-null distribution of the Wilks statistic in MANOVA. Ann. Inst. Stat. Math. **16** 67-70.

[3] Katti, S. K. (1961). Distribution of the likelihood ratio for testing multivariate linear hypothesis. Ann. Math. Statist. **32** 333-335.

[4] Khatri, C. G. (1970). A note on Mitra's paper "A density-free approach to the matrix variate beta distribution." Sankhyā, Ser. A **32** 311-318.

[5] Mitra, S. K. (1970). A density-free approach to the matrix variate beta distribution. Sankhyā, Ser. A **32** 81-88.

[6] Olkin, I. and Rubin, H. (1964). Multivariate beta distribution and independence properties of the Wishart distribution. Ann. Math. Statist. **35** 261-269.

[7] Rao, C. R. (1963). Linear Statistical Inference and Its Applications. John Wiley and Sons, New York.